AD NO. ____________________________
DTC PROJECT NO. 8-CO-160-UXO-016
REPORT NO. ATC-9364

SHALLOW WATER

UXO TECHNOLOGY DEMONSTRATION SITE

SCORING RECORD NO. 6

SITE LOCATION:
U.S. ARMY ABERDEEN PROVING GROUND

DEMONSTRATOR:
NAEVA GEOPHYSICS, INC.
P.O. BOX 7325
CHARLOTTESVILLE, VA 22906

TECHNOLOGY TYPE/PLATFORM
GEONICS EM61 HIGH POWER
(EM61 HP) SYSTEM

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

MARCH 2007

Prepared for:
U.S. ARMY ENVIRONMENTAL COMMAND
ABERDEEN PROVING GROUND, MD 21010-5401

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISPOSITION INSTRUCTIONS

Destroy this document when no longer needed. Do not return to the originator.

CSTE-DTC-AT-SL-E

MEMORANDUM FOR RECORD

SUBJECT: Operations Security (OPSEC) Review of Paper/Presentation

1. The attached record entitled "The Shallow Water UXO technology Demonstration Site Scoring Record #6" dated March 2007 is provided for review for public disclosure in accordance with AR 530-1 as supplemented. The scoring record is proposed for public release via the internet.

2. I, the undersigned, am aware of the intelligence interest in open source publications and in the subject matter of the information I have reviewed for intelligence purposes. I certify that I have sufficient technical expertise in the subject matter of this report and that, to the best of my knowledge, the net benefit of this public release outweighs the potential damage to the essential secrecy of all related ATC, DTC, ATEC, Army or other DOD programs of which I am aware.

NAME (Printed)	SIGNATURE	DATE
Gary Rowe		March 2007

CONCURRENCE:	NAME (Printed)	SIGNATURE	DATE
Program Mgr/Customer (If not ATC owned technology)	Patrick McDonnell		
Directorate Director	Charles Valz		4/11/07
Directorate OPSEC QC and Team Leader	William Burch and Michael Karwatka		27MAR07 9 APR07
ATC OPSEC Officer/ For Security Manager	Donald Lawson		12 Apr 2007
Public Affairs Specialist	Heather Hilton		20 Apr 07
Technical Director, ATC	John R. Wallace		20 Apr 07
(Return to ATC PAO for further processing)			
DTC GO/SES	N/A		

Encl
as

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
March 2007	Final	16 through 19 October 2006

4. TITLE AND SUBTITLE SHALLOW WATER UXO TECHNOLOGY DEMONSTRATION SITE SCORING RECORD NO. 6 (NAEVA/3DGEOPHYSICS, EM61 HP)	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Rowe, Gary	5d. PROJECT NUMBER 8-CO-160-UXO-016 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-AT-SL-E Aberdeen Proving Ground, MD 21005-5059	8. PERFORMING ORGANIZATION REPORT NUMBER ATC-9364
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Commander U.S. Army Environmental Command ATTN: IMAE-ATT Aberdeen Proving Ground, MD 21010-5401	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S) Same as item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES
None.

14. ABSTRACT

This report documents the efforts of NAEVA/3DGeophysics to detect and discriminate inert unexploded ordnance (UXO) using a Geonics EM61 High Power (EM61 HP) system. Testing was conducted at ATC, Standardized Shallow Water UXO Technology Demonstration Site. A description of the tested system and an estimate of survey costs along with the analysis of the system performance are provided.

15. SUBJECT TERMS

NAEVA/3DGeophysics, UXO Standardized Technology Demonstration Site, Shallow Water, EM61 HP, MEC

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified	SAR		

ACKNOWLEDGMENTS

Author:

Gary W. Rowe
Military Environmental Technology Demonstration Center (METDC)
U.S. Army Aberdeen Test Center (ATC)
U.S. Army Aberdeen Proving Ground (APG)

Contributors:

William Burch
Military Environmental Technology Demonstration Center
U.S. Army Aberdeen Test Center
U.S. Army Aberdeen Proving Ground

Christina McClung
Aberdeen Data Services Team (ADST)
Logistics Engineering and Information Technology Company (Log.Sec/Tri-S)
U.S. Army Aberdeen Proving Ground

TABLE OF CONTENTS

SECTION 4. TECHNICAL PERFORMANCE RESULTS

PAGE

SECTION 5. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of munitions and explosives of concern (MEC), i.e., unexploded ordnance (UXO) and discarded military munitions (DMM), require testing so their performance can be characterized. To that end, the U.S. Army Aberdeen Test Center (ATC) located at Aberdeen Proving Ground (APG), Maryland, has developed a Standardized Shallow Water Test Site. This site provides a controlled environment containing varying water depths, multiple types of ordnance and clutter items, as well as navigational and detection challenges. Testing at this site is independently administered and analyzed by the government for the purposes of characterizing technologies, tracking performance during system development, and comparing the performance and costs of different systems.

The Standardized UXO Technology Demonstration Site Program is a multiagency program spearheaded by the U.S. Army Environmental Command (USAEC). ATC and the U.S. Army Corps of Engineers Engineering, Research and Development Center (ERDC) provide programmatic support. The Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP), and the Army Environmental Quality Technology Program (EQT) provided funding and support for this program.

1.2 OBJECTIVE

The objective of the Shallow Water Standardized UXO Technology Demonstration Site is to evaluate the detection and discrimination capabilities of existing and emerging technologies and systems in a shallow water environment. Specifically:

a. To determine the demonstrator’s ability to survey a shallow water area, analyze the survey data, and provide a prioritized “Target List” with associated confidence levels in a timely manner.

b. To determine both the detection and discrimination effectiveness under realistic scenarios that varies ordnance, clutter, and bathymetric conditions.

c. To determine cost, time, and manpower requirements needed to operate the technology.

1.3 CRITERIA

The scoring criteria specified in the Environmental Quality Technology - Operational Requirements Document (EQT-ORD) (app D, ref 1) for: A(1.6.a): UXO Screening, Detection and Discrimination document are presented in Table 1-1. Very little information was available on the capabilities of shallow water detection systems when these criteria were developed. However, they were used in the design of the test site, and the five metrics were used to measure system performance in this report.

TABLE 1-1. SCORING CRITERIA

Metric	Threshold	Objective
Detection	80% ordnance items buried to 1 foot and under 8 feet (2.4 m) of water at a standardized site detected	95% ordnance items buried to 4 feet and under 8 feet (2.4 m) of water at a standardized site detected
Discrimination	Rejection rate of 50% of emplaced non-UXO clutter at a standardized site with a maximum false negative rate of 10%	Rejection rate of 90% of emplaced non-UXO clutter at a standardized site with a maximum false negative rate of 0.5%
Reacquisition	Reacquire within 1 meter	Reacquire within 0.5 meter
Cost rate	$4000 per acre	$2000 per acre
Production rate	5 acres per day	50 acres per day

The ATC shallow water site is designed to evaluate the threshold-detection level of a range of ordnance at the 1-foot + 8-foot requirement. Limited information is available at the objective detection level. All other measured results in this test were evaluated against both criteria levels.

1.4 APG SHALLOW WATER SITE INFORMATION

1.4.1 Location

The Aberdeen Area of APG is located in the northeast portion of Maryland on the western shore of the Chesapeake Bay in Harford County. The Shallow Water Test Site is located within a controlled range area of APG.

1.4.2 Soil Type

The area chosen for the shallow water test site was known as Cell No. 3 in a dredge-spoil field. The cell bottom is composed primarily of sediment removed from the Bush River. This is a freshwater site.

1.4.3 Test Areas

a. The test site contains five areas: calibration grid, blind test grid, littoral, open water, and deeper water. Additional detail on each area is presented in Table 1-2. A schematic of the calibration lanes is shown in Figure 1.

TABLE 1-2. TEST AREAS

Area	Description
Calibration grid	The calibration area contains 15 projectiles, 3 each 40, 60, 81, 105, and 155 mm. One of each projectile type is buried at the projectile diameter to depth ratio shown in Figure 1. This area is designed to provide the user with a sensor library of detection responses for the emplaced targets and an understanding of their resistivity prior to entering the blind test fields. Two "clutter-cloud" target scenarios have been constructed adjacent to this area (fig. 1).
Blind grid	The blind grid contains 644 detection opportunities. Each grid cell is 2 by 2 m^2. At the center of each cell is either an ordnance item, clutter, or nothing. Surrounding the blind grid on three sides are 3.6-kg (8-lb) shot puts, buried 0.3 meter deep in the sediment. The shot puts can be used as a navigational/ Global Positioning System (GPS) check. The GPS coordinates for the center of each grid and the shot put locations are provided to the vendor prior to testing.
Littoral	This is a sloping area on one side of the pond with vegetation growing into the water line. Water depth ranges from 0.3 to 1.8 meters. It contains a variety of navigational and detection challenges.
Open water	The open water scenario contains a variety of navigational, detection, and discrimination challenges. Water depth varies from 1.8 to 3.4 meters.
Deeper water	The water depth in this area varies between 3.4 and 4.3 meters.

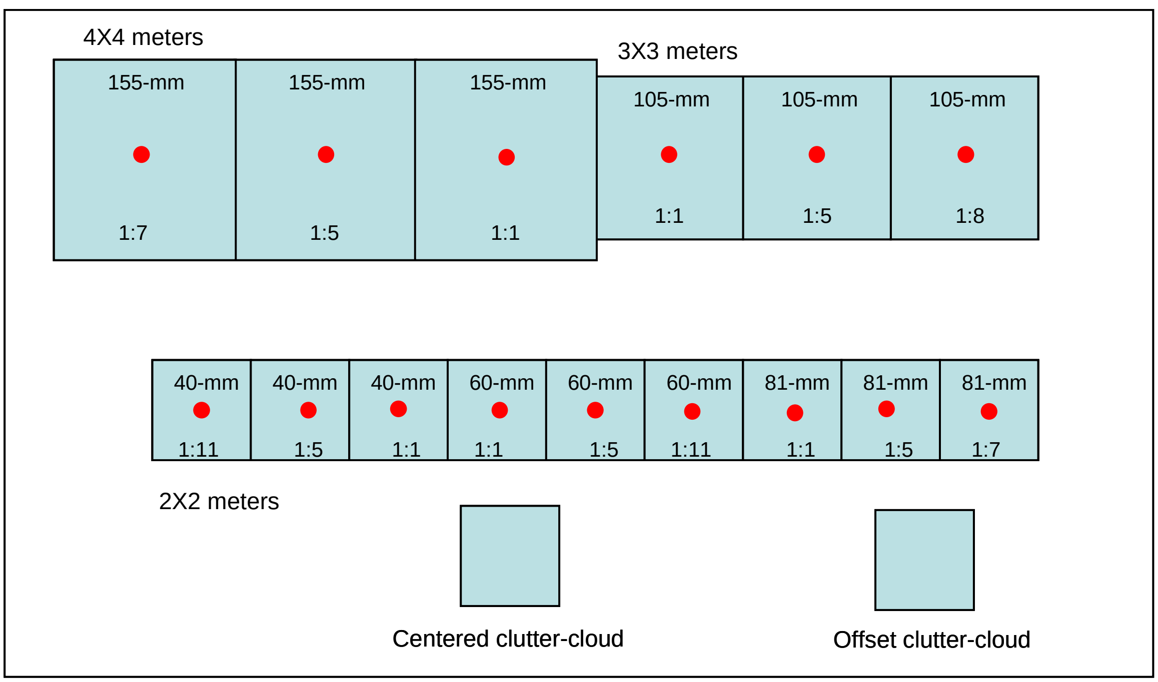

Figure 1. Schematic of the calibration grid.

b. The water depth at this facility during testing is maintained such that the calibration and blind grid areas meet the 2.4-meter (8-ft) detection criterion specified in section 1.3. The test site is approximately 2.8 hectares (6.9 acres) in size.

1.5 GROUND TRUTH TARGETS

The ground truth is composed of both inert ordnance and clutter items. The inert ordnance items are listed in Table 1-3. All items were located in storage sites at APG. The items have not been fired or degaussed.

Clutter items fit into one of three categories: ferrous, nonferrous, and mixed metals. The ferrous and nonferrous items have been further divided into three weight zones as presented in Table 1-4 and distributed throughout all test areas. Most of this clutter is composed of ordnance components; however, industrial scrap metal and cultural items are present as well. The mixed-metals clutter is composed of scrap ordnance items or fragments that have both a ferrous and nonferrous component and could reasonably be encountered in a range area. The mixed-metals clutter was placed in the open water area only.

TABLE 1-3. INERT ORDNANCE TARGETS

Description	Length, mm	Diameter, mm	Aspect Ratio, W/L	Weight, g
40-mm L70 projectile	208	40	0.1923	965
60-mm mortar M49A2	185	60	0.3243	975
81-mm mortar M374	528	81	0.1534	3969
81-mm mortar M821	510	81	0.1588	3338
105-mm projectile M1	445	105	0.2360	13834
155-mm M107 projectile	684	155	0.2266	41731
8-in. M104/106	856	203	0.2371	89811

L = Length.
W = Width.

TABLE 1-4. CLUTTER WEIGHT RANGES

Clutter Type	Weight Range in Grams		
	Small	Medium	Large
Ferrous	10 to 510	511 to 2200	> 2201
Nonferrous	10 to 270	275 to 800	> 801

SECTION 2. SYSTEM UNDER TEST

2.1 DEMONSTRATOR INFORMATION

NAEVA in conjunction with 3DGeophysics provided the information in sections 2.1 through 2.6 as part of their Broad Agency Announcement (BAA) proposal (ref 2). This information was edited to change verb tense and to conform to government report guidelines. Section 2.8 contains ATC's comments on the demonstrated system.

Note: The provided demonstrator information has been edited to comply with government report guidelines.

2.2 SYSTEM DESCRIPTION

a. The Geonics EM61 High Power (EM61 HP) system offers several advantages over the standard EM61 system. The high power system uses approximately 300 watts of transmit power instead of the approximately 100 watts in the standard system. The transmit waveform is bipolar instead of monopolar (current is driven one way and then the other and stacked). In addition, the transmit frequency of the high powered transmitter is doubled when compared with a standard system. The net result of these improvements is to increase the transmitter moment from about 150 to 1200 amperes per square meter. Thus, the signal is increased, improving the signal-to-noise ratio of the recorded data. The effect is to almost double the recordable signal from any given target at a detectable depth. This also increases the depth of penetration of the system.

b. The 3DGeophysics design for the underwater system incorporates three EM61 HP receiver coils and a single transmitter coil mounted on a carrying trailer made of thin but rugged plastic sheets with structural separators and small stainless steel bolts (see fig. 2 and 3). This design provides structural strength (a plastic sandwich) with coil pockets designed to carry receiver coils similar in form to the top coil of a standard EM61. The design incorporates a simple skid between the wheels on the undercarriage of the cart to allow the trailer to skid over rough terrain or simply wheel over even ground.

c. The proposed trailer design is configured to accept three submersible receiver coils (1 by 0.5 m) in custom-built slots to carry the coils securely. The design could also allow for two additional receiver coils to be mounted on the left and right wing of the trailer to accommodate a five-coil system. The top layer of the trailer will be configured to accept and secure a flexible transmitter loop. The entire system is designed to be semiflexible to withstand the extraordinary types of abuse that are common in towed-array work.

d. The system uses three complete Geonics EM61 consoles in one suitable field enclosure with one power connector. The system is designed and constructed to work on land as well as submerged under as much as 15 feet of water.

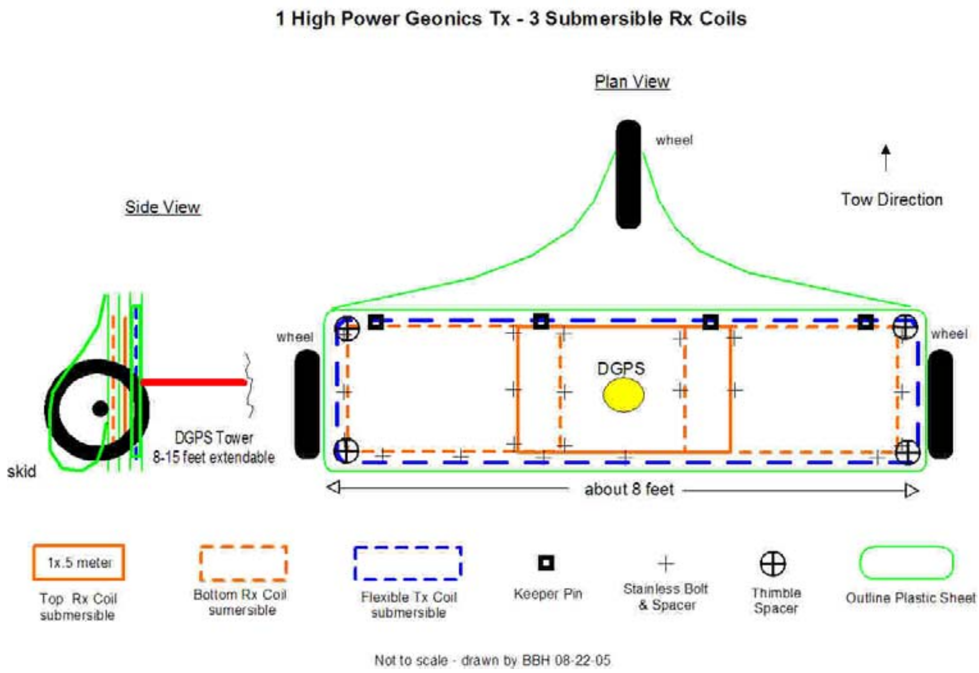

Figure 2. Envisioned schematic of the 3DGeophysics sled with EM61 coils.

Figure 3. 3DGeophysics sled in the as-tested configuration.

2.3 DEMONSTRATOR'S POINT OF CONTACT (POC) AND ADDRESS

POC: Mr. Alexander Z. Kostera
email: akostera@naevageophysics.com

Address: NAEVA Geophysics, Inc.
P.O. Box 7325
Charlottesville, VA 22906

2.4 DEMONSTRATOR'S SITE SURVEY METHOD

For this demonstration, NAEVA and 3DGeophysics proposed to deploy their multireceiver underwater system using Geonics EM61 High Power (underwater coils) electromagnetic metal detectors (one transmitter coil, three receiver coils). The system was designed to be extremely lightweight and require a small fiberglass boat for towing. This configuration allowed the team to achieve full coverage of the site, even in relatively shallow areas (fig. 4). Accurate data positioning was achieved using a Differential Global Positioning System (DGPS).

Figure 4. 3DGeophysics personnel surveying in the littoral zone.

2.5 DEMONSTRATOR'S QUALITY CONTROL (QC) AND QUALITY ASSURANCE (QA)

a. For purposes of this proposal, QA is defined as the procedures to be used during the demonstration. All of the procedures are designed to provide excellent data quality while maximizing production during the field effort.

b. All geophysical data were collected with real-time DGPS data positioning from an antenna mounted above the EMI coils. Electromagnetic data were collected at a rate of 10 readings per second, which equates to more than one reading per foot. DGPS locations were

logged at a rate of one reading every 5 seconds. To maintain straight-line profiling and to minimize the occurrence of gaps within the data, real-time sensor-tracking software was used. Positional data supplied for the calibration lanes and blind grid area are overlaid on the track map to ensure that full site coverage has been achieved. Although the DGPS has a listed accuracy of less than 10 cm, the expected accuracy of resultant target selections was signified by a circle with a 1-foot radius around each target.

c. To establish confidence in the data reliability, tests were conducted in a systematic manner throughout the duration of the fieldwork. Various types of QC data were generated before, during, and after all data collection sessions.

d. Daily: A location was identified each day that had no subsurface metal and was designated as a calibration point. Readings were collected in a stationary position over the calibration point to ensure that a stable and repeatable response was exhibited. This test was performed twice daily to establish that the instrument was functioning properly, as indicated by a stable and repeatable response.

e. A line containing at least one seeded item was identified within the calibration lanes that served as a standard response and latency check. At the start and end of each field day, two lines were collected bidirectionally across the item using, as close as possible, the same line path. The data were reviewed for consistent response and positioning and for determination of an appropriate latency correction.

f. During data collection: On completion of the original collection of a data set, approximately 5 percent of the line footage for each surveyed area was re-collected as a check of instrument repeatability and positioning. The repeat lines were saved to separate files and used to create profiles that provided direct comparison with the original data. Each profile was evaluated for repeatability in both instrument response and data positioning.

2.6 DATA PROCESSING DESCRIPTION

a. The geophysical data were temporarily stored in the system's integrated logger during data collection and downloaded into a laptop computer for on-site review and editing. Using Geosoft's Oasis Montaj software, a track plot of the instrument's DGPS positions was created to ensure that adequate data coverage had been achieved. Preliminary contour maps were created for field review of each survey area. Once in-field processing and review were completed, the data were electronically transferred to NAEVA's Virginia office for analysis and target selection.

b. Geosoft's Oasis Montaj UXO software package was used to post-process and contour the raw data and to identify potential UXO targets. The program identifies peak amplitude responses of the frequency associated with, but not limited to, UXO items. Anomalies may generate multiple target designations depending on individual signature characteristics.

c. Geophysical data processing included the following:

(1) Instrument drift correction (leveling).

(2) Lag correction.

(3) Digital filtering and enhancement (if necessary).

(4) Gridding of data.

(5) Selection of all anomalies.

(6) Selection of targets for intrusive characterization.

(7) Preparation of geophysical and target maps.

d. Final target lists for the three scenarios will be prepared separately in the specified formats and then submitted for scoring.

2.7 DEMONSTRATOR'S SITE PERSONNEL

NAEVA Project Geophysicist: Mr. Mark Howard

3DGeophysics Project Geophysicist: Mr. Brian Herridge

2.8 ATC'S SURVEY COMMENTS

a. The towing vessel used both an outboard motor at the stern of the boat and two trolling motors mounted to the port and starboard sides near the bow for propulsion and maneuvering. The outboard motor provided the power needed to tow the sled along the bottom of the pond while the thrust produced by the trolling motors helped to maneuver the boat into position for the next survey line. The trolling motors also helped counteract some of the wind and wave actions that would otherwise force the boat off the required survey heading. Experimenting using both the forward and reverse thrust from just one trolling motor led to the elimination of the second unit.

b. The design of the bottom-riding sled allows it to maneuver easily along the contours that form the shoreline and in the open water at the center of the pond. The sled rests on three swivel wheels and connects to the boat by means of a rigid pole. The combination of motors on the towing vessel, the rigid pole, and swivel wheels allows the sled to make pivot turns. Aerodynamic design elements incorporated into the plastic "sandwich" body add to the stability and towing ability of the sled in water.

c. Overall, the design of this system is highly maneuverable in a shallow water environment.

SECTION 3. SURVEY COST ANALYSIS

3.1 DATES OF SURVEY

The NAEVA/3DGeophysics electromagnetic system was tested from 16 through 19 September 2006.

3.2 SITE CONDITIONS

3.2.1 Atmospheric Conditions

An ATC weather station located adjacent to the test site recorded the average temperature and precipitation on an hourly basis for each day of operation. The temperatures listed in Table 3-1 represent the average temperature from 0700 through 1700. The hourly weather logs used to generate this summary are provided in Appendix A.

3.2.2 Water Conditions

Water conditions were monitored using a TIDALITE IV Portable Tide Gauge System®. Data recorded included water depth and temperature, significant wave height based on the average one-third wave height seen over the test period using the Draper/Tucker analysis method, and the full-wave frequency calculated by full-wave mean crossing detection. The values displayed in Table 3-1 were averaged from 0700 through 1700. Detailed information is provided in Appendix A.

TABLE 3-1. SITE CONDITION SUMMARY

Date, 06	**Air Temperature, °C**	**Wind, km/hr**	**Water Temperature, °C**	**Water Depth, m**[a]	**Significant Wave Height, m**	**Wave Frequency, Hz**
16 Oct	13	6	15	-0.2	Lost	Lost
17 Oct	14	8	15	-0.2	Lost	Lost
18 Oct	20	5	15	-0.2	Lost	Lost
19 Oct	18	5	15	-0.2	Lost	Lost

[a]Variance between the required 2.4-meter test depth and actual test conditions

Lost = Instrumentation malfunction.

3.3 SURVEY ACTIVITIES

The information contained in this section provides an estimate of the time needed and costs associated with surveying an area with this demonstrator's system. This includes data on equipment setup and calibration, site survey and any resurvey time, and downtime due to system malfunctions and maintenance requirements.

3.3.1 Survey Times

a. A government representative monitored and recorded all on-site activities, which were grouped into one of eleven categories. The first eight categories were chargeable to the system while the last three were not. Categorizing these activities provided insight into the technical and logistical aspects of the system. The times recorded in each category were then matched with the number of demonstrator personnel, assigned skill levels, and a consistent (across-vendor) salary to produce an estimate of the survey costs.

(1) Initial setup/mobilization. Started at the time the demonstrator's equipment arrived at the survey site and stopped when the system was ready to acquire data.

(2) Daily setup/close-up. Monitored time spent mounting and dismounting the equipment each day.

(3) Instrument calibration. Recorded the amount of time used for daily quality assurance checks (e.g., sensors, GPS data, survey data quality).

(4) Data collection. Time spent surveying the test area.

(5) Downtime (nonsurvey time) for equipment/data checks. Covered time spent troubleshooting equipment or verifying survey tracks.

(6) Downtime (nonsurvey time) for equipment failure. Examples include replacing damaged cables, lost communication with base station, and any other failure that prevented surveying. Some weather-related failures fall into this category, for example, light-emitting diode (LED) displays darkened by the sun and wind creating waves too high to permit surveying.

(7) Downtime (nonsurvey time) for maintenance. Battery replacement and memory downloads are typical examples.

(8) Demobilization. Commenced once the demonstrator completed the survey and concluded the final on-site check of the test data and ended when the equipment and personnel were ready to leave the site.

(9) Nonchargeable downtime for breaks and lunch. The demonstrator's company policy set this standard.

(10) Nonchargeable downtime for weather-related causes (i.e., lighting, high wet-bulb heat index, and similar events).

(11) Nonchargeable downtime due to ATC range operating requirements. Danger zone conflicts, lack of support personnel, equipment or other ATC-caused delays.

b. The daily log sheets are provided in Appendix B. Information that provides insight into the operational, maintenance, and logistic aspects of the system is summarized in Table 3-2.

TABLE 3-2. TIME ON-SITE

Date, 06	**16 Oct**	**17 Oct**	**18 Oct**	**19 Oct**	**Activity Totals, hr**
Activity (daily times recorded in minutes)					
Initial setup	180	-	-	-	3.0
Daily setup/close-up	35	120	85	55	4.9
Instrumentation calibration	-	105	10	40	2.6
Data collection	-	220	175	285	11.3
Equipment/data checks	-	-	-	-	0.0
Equipment failure	-	60	180	60	5.0
Maintenance	-	-	-	-	0.0
Demobilization	-	-	-	85	1.4
Breaks/lunch	15	-	-	-	0.3
Weather-related	-	10	-	-	0.2
ATC downtime	-	-	-	-	0.0
Daily total, hr	3.8	8.6	7.5	8.8	-

Note: Task times rounded to 5-minute increments.

3.3.2 <u>On-Site Data Collection Costs</u>

The times associated with the 11 activities have been grouped into the three basic components of the evaluation: initial setup, site survey, and pack-up (demobilization). Note that site survey time includes daily setup/stop time, data collection, breaks/lunch, downtime for equipment/data checks or maintenance, downtime due to failure, and downtime due to weather. This combines the actual survey cost with the demonstrator's associated on-site overhead costs.

A standardized estimate for labor costs associated with this effort was then calculated using the following job categories: supervisor ($95.00/hr), data analyst ($57.00/hr), and site support ($28.50/hr). The estimated costs are presented in Table 3-3.

TABLE 3-3. CALCULATED SURVEY COSTS

	No. of Persons	Hourly Wage	Hours	Cost
Initial Setup				
Supervisor	1	$95.00	3.0	$285.00
Data analyst	1	$57.00	3.0	$171.00
Site support	2	$28.50	3.0	$171.00
Subtotal				**$627.00**
Site Survey				
Supervisor	1	$95.00	24.3	$2308.50
Data analyst	1	$57.00	24.3	$1385.10
Site support	2	$28.50	24.3	$1385.10
Subtotal				**$5078.70**
Demobilization				
Supervisor	1	$95.00	1.4	$133.00
Data analyst	1	$57.00	1.4	$79.80
Site support	2	$28.50	1.4	$79.80
Subtotal				**$292.60**
Total on-site costs				**$5,998.30**

3.4 COST ANALYSIS

The data collection process described above provided an on-site cost guide to compare the performance of this vendor with any other that has demonstrated at the shallow water site. It is not a true indicator of survey costs. Many other expenses have not been included, such as travel costs, per diem, off-site data processing and analysis, company overhead, and profit.

Calculating the area surveyed was done by plotting the raw GPS coordinates and then combining the sensor swath (line spacing and associated overlap).

To determine the number of acres surveyed per day, the total number of hours spent at the test site (table 3-2) was divided by 8 (converts to 8-hr days). The number of acres was then divided by the number of 8-hour days. The cost per acre was determined by dividing the total survey costs (table 3-3) by the same number of acres. This information is summarized in Table 3-4.

TABLE 3-4. SURVEY COSTS

Area surveyed (acre[a])	2.8
Time on-site (8-hr days)	4.15
Calculated survey cost (U.S. dollars)	$5998.30
Acres per day	0.67
Cost per acre	$2142.25

[a]Acre = 4047 m^2.

A comparison of the NAEVA/3DGeophysics survey costs with the EQT-ORD criteria is presented in Table 3-5.

TABLE 3-5. TEST RESULTS - CRITERIA COMPARISON

Metric	Threshold	Objective	NAEVA/3DGeophysics
Cost rate	$4000 per acre	$2000 per acre	$2142.25
Production rate	5 acres per day	50 acres per day	0.67

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 AREA SURVEYED

4.1.1 Calculated Area

a. Both the test and scoring methodologies required the demonstrator to survey 100 percent of each of the four test areas (blind grid, open water, littoral, and deeper water). Scoring a partially surveyed area alters the ordnance and clutter sample sizes, and test area boundaries, and decreases the statistical confidence in the performance statements made for that area. Allowing partial scoring decreases the validity of performance comparisons made between multiple test areas for a single demonstrator and comparisons made between multiple demonstrators for a single test area.

b. Realizing that some systems may not be able to survey 100 percent of a given test area, a ranking system was established. The percent coverage for a given test area is determined by first plotting the raw GPS coordinates combined with the sensor swath (line spacing and associated overlap), calculating the area surveyed, and then comparing the surveyed area with the total test area.

$$\frac{\text{Section Surveyed}}{\text{Test Area Size}} \times 100 = \% \text{ Surveyed}$$

c. The demonstrator's system is always scored against the complete ground truth for a given test area regardless of the percentage covered.

4.1.2 Area Assessment

The ranking system and survey results are presented in Table 4-1.

TABLE 4-1. M882 SURVEY RANKING SYSTEM AND RESULTS

Ranking System		Survey Results		
% Area Covered	**Ranking**	**Test Area**	**% Area Covered**	**Data Use**
95 to 100	Met	Blind grid	100	Direct comparison between systems and areas.
90 to 94	Generally met	-	-	Comparison between systems and areas. A small negative bias is contained in the reported numbers (bias not quantified in this report).
50 to 89	Partially met	Littoral	84	Reported, not compared between systems or areas. A large negative bias is contained in the reported numbers (bias not quantified in this report).
0 to 49	Not met	-	-	Not scored/not reported.

Two of the four test areas within the shallow water site were damaged during a prior demonstration. An undetermined percentage of projectiles in the open and deeper water areas, that were either pressed flush with or resting on top of the pond bottom, have been dislodged and dragged out of their original locations. Accurately measuring system performance in these areas is not possible. The scope of this demonstration was reduced to the blind grid and littoral test areas only.

4.2 SYSTEM SCORING PROCEDURES

a. The scoring entities used in this program were predicated on knowing the composition and location of every detectable item in an area. The deeper water area is the one exception. Ground truth targets were placed in this area without a pre-survey and clearing operation. Therefore, only the system's probability of detection (P_d) was evaluated in this area.

b. The best indicator of survey performance is the blind grid. This area provides a statically valid, controlled environment in which the demonstrator must provide a response (ordnance, clutter, or blank) at each of the 644 locations. Comparison of the response and discrimination lists to the ground truth in this area both determines the range of ordnance the system can reliably detect and establishes the baseline to which system performance in all other test areas is measured.

c. The scoring terms and definitions, along with an explanation of the receiver operating characteristics (ROC) curve development and the chi-square analysis used in this report, are provided in Appendix C.

d. Demonstrator performance was scored in two stages: response and discrimination.

e. Response stage scoring evaluates the ability of the demonstrator's system to detect emplaced ground truth targets without regard to discriminating ordnance from clutter. In this stage, the GPS locations and signal strengths of all anomalies the demonstrator deemed sufficient for further investigation and/or processing are reported. This list was generated with minimal processing, i.e., associating signal strength with GPS location, and includes only signals that are above the system noise level.

f. The discrimination stage evaluated the demonstrator's ability to segregate ordnance from clutter. The same GPS locations reported in the response stage anomaly list were evaluated on the basis of the demonstrator's discrimination process (section 2.6). A discrimination stage list was generated and prioritized on the basis of the demonstrator's determination that an anomaly was more likely to be ordnance rather than clutter. Typically, higher output values indicate a higher confidence that an ordnance item is present at a specified location. The demonstrator then specifies the threshold value for the prioritized ranking that provides optimal system performance. This value is the discrimination stage threshold.

g. Both the response and discrimination lists contain the identical number of potential target locations, differing only in the priority ranking of the declarations.

h. Within both of these stages, the following entities were measured:

(1) P_d.

(2) Probability of false positive (P_{fp}).

(3) Probability of background alarm (P_{ba})/background alarm rate (BAR).

4.2.1 ROC Curves

a. Based on the entire range of ground truth targets used at this site, ROC curves were generated for both the response and discrimination stages. In both stages, the probability of detection versus false alarm rates was plotted. False alarms were divided into two groups: (1) anomalies corresponding to emplaced clutter items, thereby measuring the P_{fp} and (2) anomalies not corresponding to any known item, termed background alarms (P_{ba}) in the blind grid area and BAR in all other areas.

b. The ROC curves for the response and discrimination stages for all areas surveyed are shown in Figures 5 through 8. Horizontal lines illustrate the system performance at the demonstrator's recommended noise level during the response stage, or discrimination threshold level in the discrimination stage. The point where the curve crosses the horizontal line defines the subset of targets the demonstrator recommends digging.

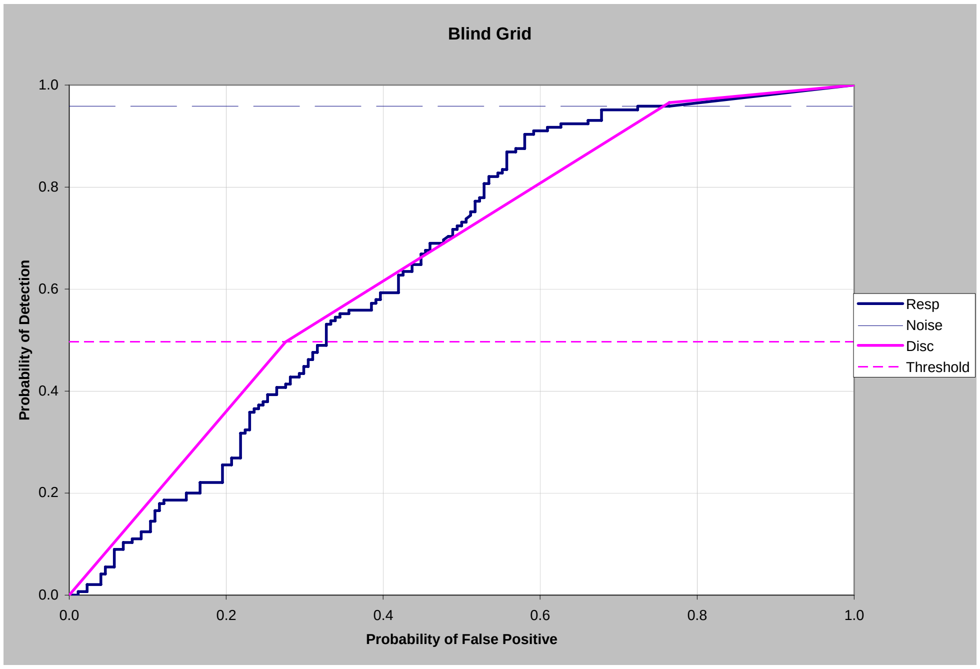

Figure 5. Blind grid P_d versus P_{fp}.

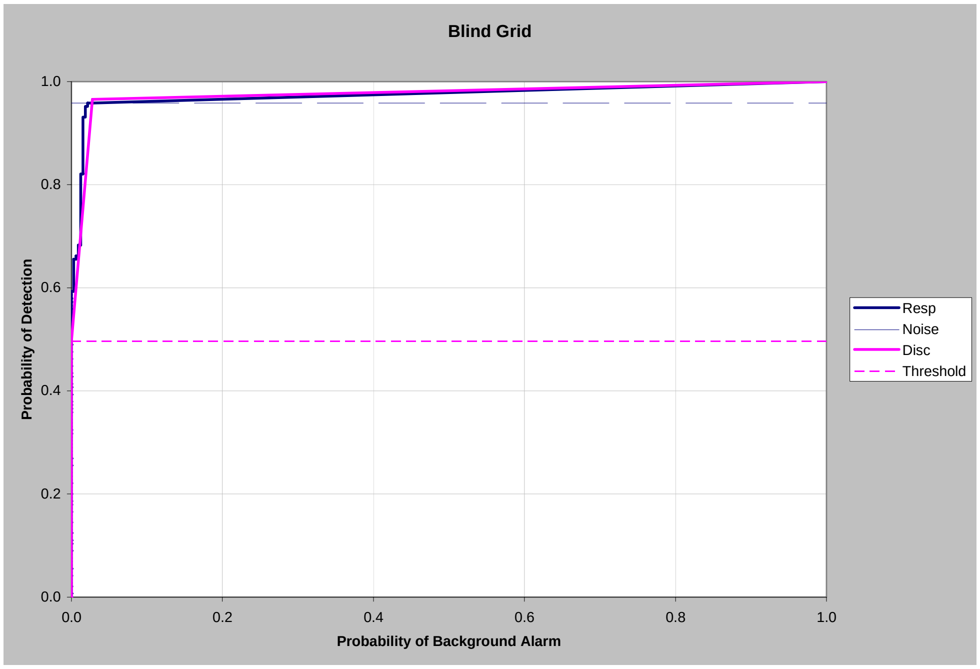

Figure 6. Blind grid P_d versus P_{ba}.

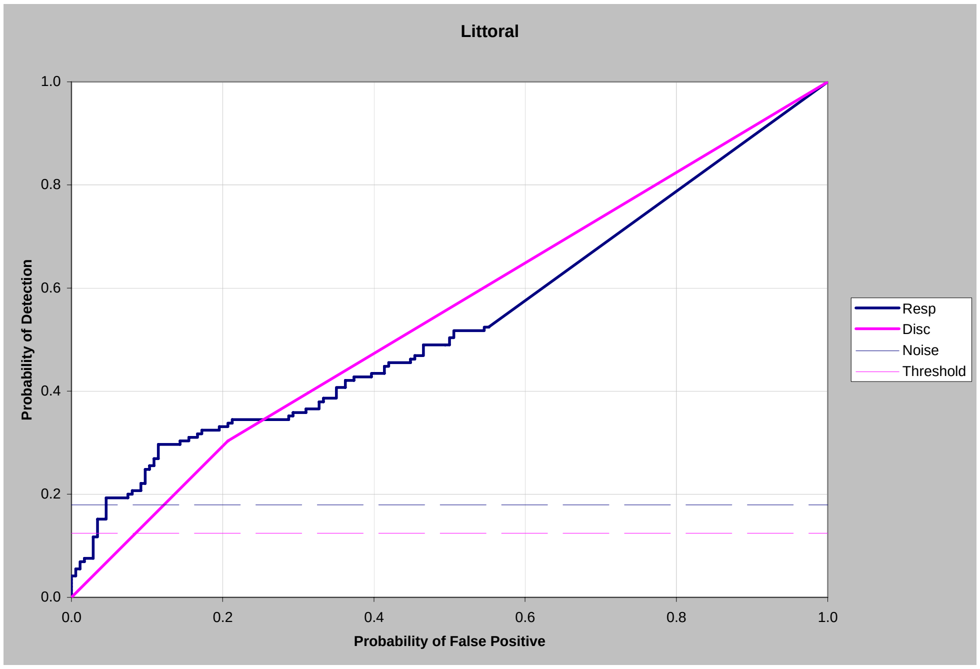

Figure 7. Littoral P_d versus P_{fp}.

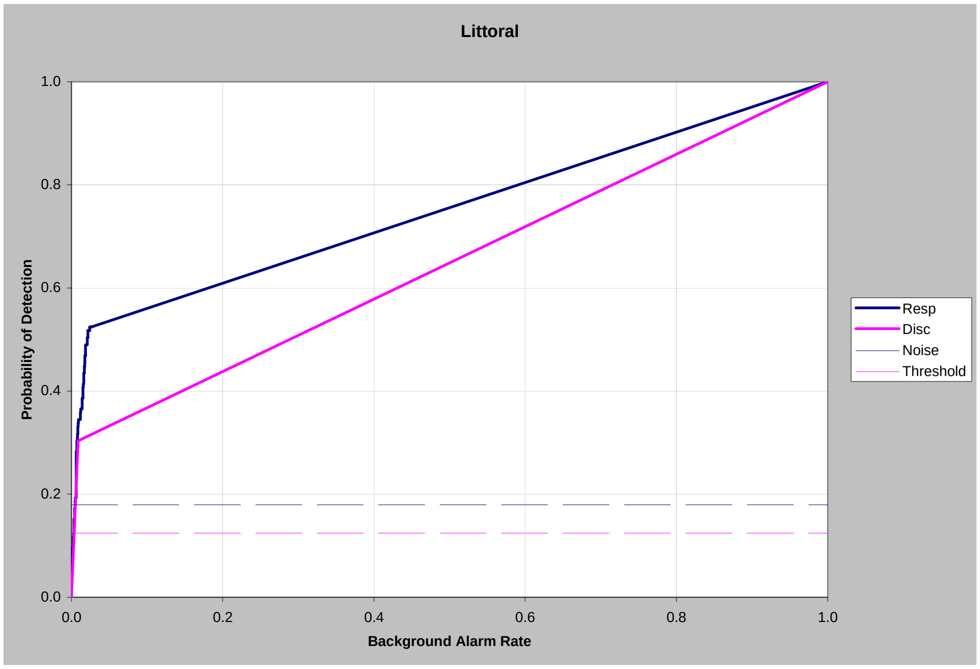

Figure 8. Littoral P_d versus BAR.

4.2.2 Detection Results

Detection results, broken out by stage, area surveyed, and ordnance size, are presented in Table 4-2. The results by size indicate how well the demonstrator detected/discriminated ordnance of a given caliber. Overall results summarize ordnance detection over a given area. All values were calculated assuming the number of detections was a binomially distributed random variable. These results are reported at the 90-percent reliability/95-percent confidence levels unless otherwise noted.

TABLE 4-2. SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber					
		40 mm	60 mm	81 mm	105 mm	155 mm	8 in.
Blind grid							
Response stage							
P_d	95.2%	96.6%	93.1%	93.1%	100.0%	93.1%	
P_d lower 90% confidence	92.0%	87.2%	82.7%	82.7%	92.4%	82.7%	
P_{fp}	92.0%						
P_{fp} lower 90% confidence	88.6%						
P_{ba}	5.8						
Discrimination stage							
P_d	45.5%	27.6%	44.8%	37.9%	55.2%	62.1%	
P_d lower 90% confidence	39.9%	16.8%	31.9%	25.7%	41.7%	48.5%	
P_{fp}	56.9%						
P_{fp} lower 90% confidence	51.8%						
P_{ba}	1.8						
Littoral region							
Response stage							
P_d	17.9%	13.8%	6.9%	27.6%	17.2%	24.1%	
P_d lower 90% confidence	13.9%	6.2%	1.8%	16.8%	8.6%	14.0%	
P_{fp}	15.5%						
P_{fp} lower 90% confidence	12.0%						
BAR m^{-2}	0.012						
Discrimination stage							
P_d	12.4%	13.8%	0.0%	27.6%	6.9%	13.8%	
P_d lower 90% confidence	9.0%	6.2%	0.0%	16.8%	1.8%	6.2%	
P_{fp}	5.7%						
P_{fp} lower 90% confidence	3.6%						
BAR m^{-2}	0.004						
Response stage noise level: 160							
Recommended discrimination threshold: 1							

4.2.3 System Discrimination

Using the demonstrator's recommended setting, the items detected and correctly classified as ordnance were further evaluated as to whether the demonstrator could correctly identify the ordnance type. The list of ground truth ordnance items was provided to the demonstrator prior to testing.

The NAEVA/3DGeophysics "dig-list" discriminated between ordnance and clutter but not between ordnance types. The latter was an optional requirement.

4.2.4 System Effectiveness

Efficiency and rejection rates were calculated to quantify the discrimination ability at two specific points of interest on the ROC curve: the point where no decrease in P_d occurred (i.e., the efficiency is by definition equal to one) and the operator-selected threshold. These values, for both magnetometers, are presented in Table 4-3.

TABLE 4-3. EFFICIENCY AND REJECTION RATES

	Efficiency	False Positive Rejection Rate	Background Alarm Rejection Rate
Blind Grid			
At operating point	0.52	0.64	1.00
With no loss of P_d	1.00	0.64	1.00
Littoral Region			
At operating point	0.69	0.63	0.63
With no loss of P_d	1.00	0.63	0.63

4.2.5 Chi-Square Analysis

Typically, this report contains a chi-square 2-by-2 contingency test for comparison between ratios to compare performance across test areas with regard to P_d^{res}, P_d^{disc}, P_{fp}^{res}, and P_{fp}^{disc}, efficiency, and false alarm rejection rates. The intent of the comparison is to determine if the features introduced in each test region have a degrading effect on the performance of the sensor system.

This system did not survey enough of the other test areas to permit a valid comparison of performance between the areas.

4.2.6 Location Accuracy

The data points in the scatter graph shown in Figure 9 represent the coordinates of ordnance items in the littoral test area that were first detected in the response stage within a 0.5-meter radius of their true positions, then correctly identified as ordnance in the discrimination stage. The maximum error represents the 0.5-meter detection limit. The mean error represents the statistical mean of the sample considered.

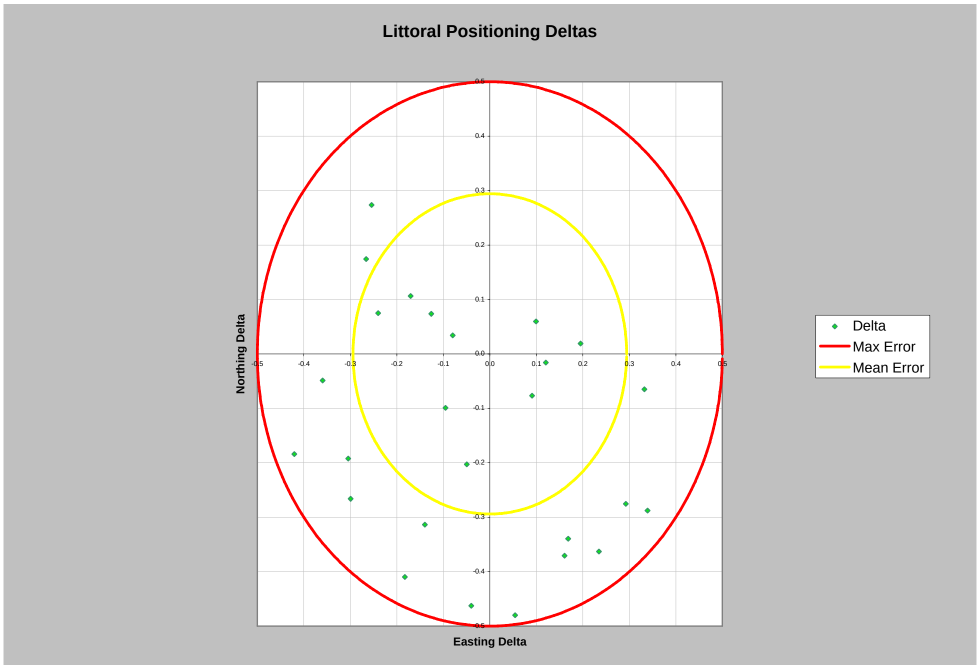

Figure 9. NAEVA/3DGeophysics littoral positioning deltas.

A visual analysis of the data point distribution shows an identical number of points (9) in quadrants III and IV, 6 in II, and 2 points in quadrant I. This suggests there may be a positioning bias in the system.

Comparisons made between the results obtained during testing and the EQT-ORD criteria are in Table 4-4.

TABLE 4-4. NAEVA/3DGEOPHYSICS TEST RESULTS - CRITERIA COMPARISON

Metric	Threshold	Objective	By Area	
Detection	80% ordnance items buried to 1 foot and under 8 feet (2.4 m) of water	95% ordnance items buried to 4 feet and under 8 feet (2.4 m) of water	Blind grid	95.9%
			Littoral	17.9%
Discrimination	Rejection rate of 50% of emplaced non-UXO clutter	Rejection rate of 90% of emplaced non-UXO clutter	Blind grid	64%
			Littoral	63%
	Maximum false negative rate of 10%	Maximum false-negative rate of 0.5%	Not assessed. An analytical procedure is not available to address this criterion.	
Reacquisition	Reacquire within 1 meter	Reacquire within 0.5 meter	The reported detection values are based on ordnance items identified within 0.5 meter of the georeferenced ground truth targets.	

Note: The blind grid and open water areas are in general accordance with the threshold requirements.

SECTION 5. APPENDIXES

APPENDIX A. TEST CONDITIONS LOG

ATMOSPHERIC CONDITIONS

Date, 06	Time, EDT	Average Wind Direction, deg	Average Wind Speed, km/hr	Wind Direction Average Standard Deviation, deg	Peak Wind Speed, km/hr	Average Temperature, °C
16 Oct	0700	184	2.5	12	5.4	3.3
	0800	51	1.4	11	4.3	4.4
	0900	48	2.9	21	6.5	9.9
	1000	86	4.3	30	9.7	12.4
	1100	183	5.0	38	10.1	14.9
	1200	205	9.4	16	14.8	15.8
	1300	203	9.4	21	16.6	16.7
	1400	206	11.2	15	18.0	16.7
	1500	220	10.4	13	15.8	16.9
	1600	198	8.6	12	12.6	16.5
	1700	180	4.0	3	6.5	15.6
17 Oct	0700	123	6.8	31	16.6	13.3
	0800	126	6.1	28	14.4	13.6
	0900	119	6.5	31	15.1	13.5
	1000	122	7.9	28	18.4	13.9
	1100	137	9.4	29	21.6	14.3
	1200	141	9.0	28	21.6	14.3
	1300	139	10.1	24	23.4	14.6
	1400	135	7.9	26	20.5	14.9
	1500	127	7.6	27	17.3	15.1
	1600	142	8.6	24	18.4	15.4
	1700	148	9.7	22	20.5	15.9
18 Oct	0700	225	0.7	2	2.2	16.7
	0800	227	3.6	7	5.4	17.2
	0900	284	3.6	17	6.1	17.6
	1000	309	2.5	27	6.1	18.2
	1100	331	3.2	21	6.8	18.8
	1200	264	5.0	35	10.8	19.9
	1300	316	4.0	23	8.3	21.6
	1400	233	7.6	21	12.6	22.1
	1500	194	10.4	7	14.0	21.4
	1600	196	8.3	10	12.2	21.9
	1700	201	5.4	5	7.9	21.6
19 Oct	0700	129	2.2	22	4.3	15.3
	0800	37	2.5	6	4.0	15.8
	0900	109	1.1	18	3.2	16.8
	1000	106	2.5	18	5.4	17.3
	1100	111	3.6	19	7.2	17.6
	1200	223	1.8	35	4.3	18.0
	1300	249	6.5	30	11.2	18.6
	1400	185	8.6	11	14.0	18.2
	1500	192	10.4	9	15.5	18.3

The TIDALITE IV Portable Tide Gauge System® was not operational. Manual water depth and temperature measurements were recorded each morning. The single measurements for each day are presented in Table 3-1.

APPENDIX B. DAILY ACTIVITIES LOG

(Page B-2 Blank)

Company: NAEVA Date: 16 October 2006			**On-site Personnel: Mark Howard, Brian Herridge, Brian Neely**	
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0915	0930	Arrived at test site. Safety briefing/question and answer session.	Downtime ATC	-
0930	1230	Mounting EM coils to sled. Could not move the boat because the correct size ball for the trailer hitch was not on site. Instrumented as much of the boat as possible on land. Set up base station.	Initial setup	180
1230	1245	Lunch.	Nonchargeable downtime	15
1245	1320	Accomplished as much as possible without putting the boat in the water.	Daily close-up	35

Company: NAEVA Date: 17 October 2006			**On-site Personnel: Mark Howard, Brian Herridge, Brian Neely, Erik Kitt**	
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0845	0955	Arrived at site. Light rain. Completed attaching coils to the sled. Launched the boat, attached the sled, instrumented the boat, and attached EM cables. Set up base station.	Daily setup	70
0955	1005	Static calibration.	Calibration	10
1005	1015	Wind and rain increasing, adjusting tarp over the boat to compensate.	Downtime equipment	10
1015	1150	Maneuvered coils into a quiet area of the pond. Coils reaching operating temperature, taking background readings.	Calibration	95
1150	1400	Surveying calibration lane and blind grid. Steady rain.	Data collection	130
1400	1500	Computer USB ports wet, not getting data.	Downtime equipment	60
1500	1630	Blind grid survey.	Data collection	90
1630	1720	Cleanup.	Daily close-up	50

Company: NAEVA Date: 18 October 2006			**On-site Personnel: Mark Howard, Brian Herridge, Brian Neely, Erik Kitt**	
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0850	0930	Setup.	Daily setup	40
0930	1125	Hardened computer used for EM data collection failed (water from yesterday's rain). Remapping computer ports on a laptop computer for use.	Downtime equipment	115
1125	1135	Static calibration.	Calibration	10
1135	1225	Surveying.	Data collection	50
1225	1255	GPS signal dropped out.	Downtime equipment	30
1255	1500	Surveying.	Data collection	125
1500	1535	Trouble shooting laptop computer crashed.	Downtime equipment	35
1535	1620	Cleanup.	Daily closeup	45

Company: NAEVA Date: 19 October 2006			**On-site Personnel: Mark Howard, Brian Herridge, Brian Neely, Erik Kitt**	
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0810	0905	Setup.	Daily setup	55
0905	0945	Nulling coils.	Calibration	40
0945	1000	Surveying.	Data collection	15
1000	1100	Returned to dock. Amplitude on one coil is creeping upward. Problem not resolved. Decided to resume survey; will decide how to use the collected data during the processing stage.	Downtime equipment	60
1100	1530	Surveying the littoral zone.	Data collection	270
1530	1655	Cleanup/packed up.	Demobilization	85

APPENDIX C. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Munitions and Explosives of Concern (MEC): Specific categories of military munitions that may pose unique explosive safety risks, including UXO as defined in 10 USC 101(e)(5), DMM as defined in 10 USC 2710(e)(2) and/or munitions constituents (e.g. TNT, RDX) as defined in 10 USC 2710(e)(3) that are present in high enough concentrations to pose an explosive hazard.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., nonordnance item) buried by the government at a specified location in the test site.

R_{halo}: A predetermined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. For the purpose of this program, a circular halo 0.5 meter in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meter in length. When ordnance items are longer than 0.6 meter, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the projected length of the ordnance onto the ground plane plus 1 meter.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the blind grid test area.

Discrimination Stage Threshold: The demonstrator selects the threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response stage detections)/(No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm: An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open water or littoral scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): blind grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open water only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can, therefore, be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to nonordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm: An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open water or littoral scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination stage signal strength. These quantities can, therefore, be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

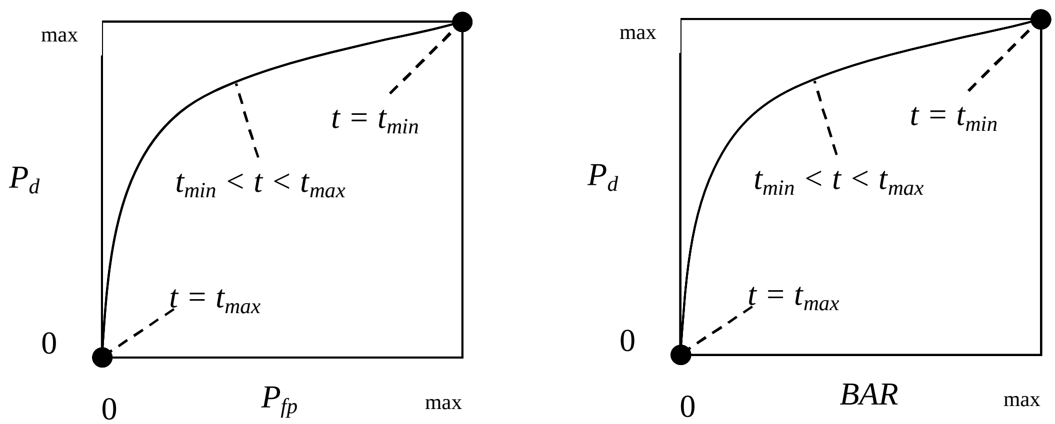

Figure A-1. ROC curves for open-site testing. Each curve applies to both the response and discrimination stages.

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a predetermined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an open water scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open water ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the blind grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from nonordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$: measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage t_{min}) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$: measures (at a threshold of interest), the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage t_{min}). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind Grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$
Open water: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})])$

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 3, pages 144 through 151).

A one-sided 2 x 2 contingency table is used in the Shallow Water Site Program to compare each area (open water, littoral, deep water) to the blind grid since each area introduces a water feature that makes it potentially more difficult to survey than the blind grid. The one-sided 2 x 2 contingency table is used to determine if there is reason to believe that the

proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly degraded by the more challenging feature introduced. A two-sided 2 x 2 contingency table is used to compare performance between any two of the test sites other than the blind grid, to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly different between those two test sites.

The test statistic of the 2 x 2 contingency table is the chi-square distribution with one degree of freedom. For the one-sided test, a significance level of 0.05 is chosen which sets a critical decision limit of 3.84 from the chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The chi-square test cannot be used in these instances. Instead, Fisher's Exact Test is used, and the critical decision limit is the chosen significance level, which is 0.05 for one-sided tests and 0.10 for two-sided tests. With Fischer's test, if the test statistic (p-value) is less than the critical value, then the null hypothesis of similar performance is rejected in favor of the alternative hypothesis: significantly greater than for the one-sided case or significantly different for the two-sided case.

Shallow Water UXO Detection Test Site examples, where blind grid results are compared to those from the open water and littoral sites and the nongrid sites (open water and littoral), are compared to each other as follows. It should be noted that a significant result does not prove a cause and effect relationship exists between the change in survey area and sensor performance; however, it does serve as a tool to indicate that one data set reflects relatively degraded system performance of a large enough scale than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind Grid	Open water	Littoral
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN WATER. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the

open water. Fisher's test must be used since a 100 percent success rate occurs in the data. Fisher's test uses the four input values to calculate a test statistic (p-value) of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open water relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN WATER. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 out of 10 emplaced ordnance items were correctly discriminated as such in open water testing. Those four values are used in the chi-square Contingency Test to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 3.84, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: BLIND GRID versus LITTORAL. Using the example data above to compare probabilities of detection in the response stage, 100 out of 100 and 20 out of 33 are used to calculate a test statistic (< 0.000) that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.61) is considered to be significantly less at the 0.05 level of significance.

P_d^{disc}: BLIND GRID versus LITTORAL. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 and 8 out of 33 emplaced ordnance items were correctly discriminated as such in open water testing. Those four values are used to calculate a test statistic of 32.01. Since the test statistic is greater than the critical value of 3.84, the smaller discrimination stage detection rate (0.24) is considered to be significantly less at the 0.05 level of significance.

P_d^{res}: OPEN WATER versus LITTORAL. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.10 level of significance.

P_d^{disc}: OPEN WATER versus LITTORAL. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the two discrimination stage detection rates are considered to be significantly different at the 0.10 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and change in performance, it does indicate that the ability of Demonstrator X to correctly discriminate seems to have been degraded by features of the littoral area relative to results from the open water using the same system.

APPENDIX D. REFERENCES

1. Environmental Quality Technology - Operational Requirements Document (EQT-ORD) for: A(1.6.a): UXO Screening, Detection and Discrimination.

2. Proposal for Shallow Water Unexploded Ordnance (UXO) Detection & Discrimination Technology Demonstration Volume 1 - Technical/Management (Plan). Submitted in response to BAA W91ZLK-04-R-0001, by NAEVA Geophysics, Inc., 30 August 2005.

3. Practical Nonparametric Statistics, W.J. Conover, John Wiley & Sons, 1980, pages 144 through 151.

APPENDIX E. ABBREVIATIONS

ADST	=	Aberdeen Data Services Team
APG	=	Aberdeen Proving Ground
ATC	=	U.S. Army Aberdeen Test Center
BAA	=	Broad Agency Announcement
BAR	=	background alarm rate
DGPS	=	Differential Global Positioning System
DMM	=	discarded military munitions
EM	=	electromagnetic
EQT	=	Army Environmental Quality Technology Program
EQT-ORD	=	Environmental Quality Technology - Operational Requirements Document
ERDC	=	U.S. Army Corps of Engineers Engineering, Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
GPS	=	Global Positioning System
LED	=	light-emitting diode
MEC	=	munitions and explosives of concern
METDC	=	Military Environmental Technology Demonstration Center
NAEVA	=	North American Exploration of Virginia, Inc.
P_{ba}	=	probability of background alarm rate
P_d	=	probability of detection
P_d^{disc}	=	probability of detection, discrimination stage
P_d^{res}	=	probability of detection, response stage
P_{fp}	=	probability of false positive
P_{fp}^{disc}	=	probability of false positive, discrimination stage
P_{fp}^{res}	=	probability of false positive, response stage
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver operating characteristics
SERDP	=	Strategic Environmental Research and Development Program
USAEC	=	U.S. Army Environmental Command
USB	=	universal serial bus
UXO	=	unexploded ordnance

APPENDIX F. DISTRIBUTION LIST

DTC Project No. 8-CO-160-UXO-016

Addressee	No. of Copies
Commander	
U.S. Army Aberdeen Test Center	
ATTN: CSTE-DTC-AT-SL-E (Mr. Gary Rowe)	1
(Library)	1
CSTE-DTC-AT-CS-R	1
Aberdeen Proving Ground, MD 21005-5059	
Commander	
U.S. Army Environmental Command	
ATTN: IMAE-ATT (Ms. Kimberly Watts)	1
Aberdeen Proving Ground, MD 21010-5401	
Booz Allen Hamilton	
ATTN: (Mr. Patrick McDonnell)	1
4692 Millennium Drive, Suite 200	
Belcamp, MD 21017-1535	
NAEVA Geophysics, Inc.	
ATTN: (Mr. Alexander Z. Kostera)	1
P.O. Box 7325	
Charlottesville, VA 22906	
Defense Technical Information Center	
8725 John J. Kingman Road, Suite 0944	PDF
Fort Belvoir, VA 22060-6218	

Secondary, "hard-copy" distribution is controlled by Commander, U.S. Army Environmental Command, ATTN: IMAE-ATT.